Federal Aviation
Administration

DOT/FAA/AM-14/6
Office of Aerospace Medicine
Washington, DC 20591

Usability of Light-Emitting Diodes in Precision Approach Path Indicator Systems by Individuals With Marginal Color Vision

Nelda J. Milburn
Kevin M. Gildea
Deborah L. Perry
Carrie A. Roberts
Linda (Sarah) Peterson
Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

May 2014

Final Report

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-14/6		

4. Title and Subtitle	5. Report Date
Usability of Light-Emitting Diodes in Precision Approach Path Indicator Systems by Individuals With Marginal Color Vision	May 2014
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Milburn NJ, Gildea KM, Perry DL, Roberts CA, Peterson LM	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125	11. Contract or Grant No.
12. Sponsoring Agency name and Address	13. Type of Report and Period Covered
Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591	14. Sponsoring Agency Code

15. Supplemental Notes

Work was accomplished under approved task AM-521.

16. Abstract

To save energy, the FAA is planning to convert from incandescent lights to light-emitting diodes (LEDs) in precision approach path indicator (PAPI) systems. Preliminary work on the usability of LEDs by color vision-waivered pilots (Bullough, Skinner, & Milburn, 2012) indicated that red weak (protan) individuals made a few errors identifying red. Hence, this follow-up study explored whether clustering LEDs of different chromaticities of the same hue would aid users with color vision deficiencies (CVD).

Participants, aged 18-33 years, included 45 with normal color vision (NCV). Additionally, 48 were diagnosed using the Colour Assessment and Diagnosis (CAD) test as 27 deutans (including five subjects with potential deutan deficiencies), 11 protans, two tritans, and eight subjects evidencing both red-green and yellow-blue deficiencies. Participants completed the Dvorine pseudoisochromatic plate test, the Signal Light Gun Test (which is the secondary test for issuing color vision waivers), the Cone Contrast Test, and a simulation of the PAPI system. Participants were asked to name the colors of a 4-light, color-coded PAPI simulation using typical red (R) and white (W) light configurations resulting in five possible patterns (WWWW, WWWR, WWRR, WRRR, and RRRR) composed of either incandescent lights, or clusters of three monochromatic or three heterochromatic 5mm cylindrical LEDs.

Analyses did not indicate any significant differences between the incandescent, monochromatic, and heterochromatic conditions. A between-group analysis found that a group comprised of those with both red-green and yellow-blue deficiencies performed significantly worse than all other color vision groups. Performance was not significantly different among all other groups. The protan group performed perfectly on all light sources, even on the 16 trials without comparative color luminance cues.

This finding suggests that color vision-waivered pilots will perform the same on red and white LED PAPI systems as with the current incandescent system.

17. Key Words	18. Distribution Statement
Color Vision Deficiency, Color Vision, Color Vision Screening Tests, Light-Emitting Diodes, Precision Approach Path Indicator	Document is available to the public through the Internet: www.faa.gov/go/oamtechreports

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	30	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

ACKNOWLEDGMENT

Research reported in this paper was conducted under the Flight Deck Program Directive / Level of Effort Agreement between the Federal Aviation Administration Headquarters and the Aerospace Human Factors Division of the Civil Aerospace Medical Institute sponsored by the Office of Aerospace Medicine and supported through the FAA NextGen Human Factors Division.

Contents

Usability of Light-Emitting Diodes in Precision Approach Path Indicator Systems by Individuals With Marginal Color Vision

DEFINITIONS AND ABBREVIATIONS

Abbreviations

ANOVA ---Analysis of Variance
AME --- ---Aviation Medical Examiner
CAD- --- ---Colour Assessment and Diagnosis Test
CAMI --- ---Civil Aerospace Medical Institute
CIE -- --- ---Commission Internationale de l'Eclairage
CCT- --- ---Cone Contrast Test
CCT (K) ---Correlated Color Temperature in Kelvin
CFR - --- ---Code of Federal Regulations
CVD --- ---Color Vision Deficient/Color Vision Deficiency
FAA-- --- ---Federal Aviation Administration
GLM --- ---General Linear Model
HC -- --- ---Heterochromatic: of, having, or consisting of different or contrasting colors; many-colored
Hetero -- ---Short for heterochromatic
ICAO --- ---International Civil Aviation Organization
L-cones - ---Long-wavelength sensitive cones
M-cones ---Medium-wavelength sensitive cones
MC -- --- ---Monochromatic: of, having, or consisting of the same color; one color
MFT --- ---Medical Flight Test
Mono --- ---Short for monochromatic
NCV --- ---Normal Color Vision
NTSB --- ---National Transportation Safety Board
OCVT -- ---Occupational Color Vision Test
PAPI - --- ---Precision Approach Path Indicator
PIP--- --- ---Pseudoisochromatic Plate
RG --- --- ---Red-Green
RPI -- --- ---Rensselaer Polytechnic Institute
S-cones-- ---Short-wavelength sensitive cones
SD --- --- ---Standard Deviation
SODA -- ---Statement of Demonstrated Ability
SLG-- --- ---Signal Light Gun
SLGT --- ---Signal Light Gun Test
SNU - --- ---Standard Normal Unit
UK--- --- ---United Kingdom
USDOT ---United States Department of Transportation
YB --- --- ---Yellow-Blue

Nomenclature

cd/m^2 --- ---Candelas per square meter
cm --- --- ---Centimeters
%- --- --- ---Percent
ft - --- --- ---Feet
K - --- --- ---Degrees Kelvin
mm -- --- ---Millimeters
nm --- --- ---Nanometers
s -- --- --- ---Second (time)
W/m^2 --- ---Watts per square meter

Usability of Light-Emitting Diodes in Precision Approach Path Indicator Systems by Individuals With Marginal Color Vision

INTRODUCTION

Cost analyses support the replacement of the current (incandescent) airport lighting system with a system using more energy-efficient light emitting diodes (LEDs); however, the usability of the colored LEDs by individuals with color vision waivers is a concern. A further concern is that the Federal Aviation Administration's (FAA's) current color vision screening tests were validated using an incandescent light source, and validity results may not directly relate to tasks involving LEDs. The incandescent source, the Signal Light Gun, is an instrument used by air traffic controllers to communicate with pilots that are experiencing radio failure. The Signal Light Gun Test (SLGT) has served, for several years, as a practical evaluation of whether an individual could distinguish the red, green, and white incandescent lights that are commonly used in airports as signal lights and often employed in safety-critical color-coded tasks such as the red and white color coding of precision approach path indicator (PAPI) lights. Furthermore, the SLGT is the instrument used to provide evidence of sufficient color vision ability for pilots to obtain a color vision waiver in the form of a Statement of Demonstrated Abilities (SODA) for aeromedical screening (CFR, 2014).

The FAA (1988) and ICAO (1988) established chromaticity polygons for aviation colors within the Commission Internationale de l'Eclairage (CIE) color space based on incandescent light presentations. Concerns from the FAA Flight Technologies and Procedures Division (AFS-400) that the narrow band, monochromatic light output of LEDs may present a different visual cue for pilots with color vision deficiencies than the broad band incandescent light sources led to this research. The purpose of the research is to provide empirical evidence that critical color-coding produced by LED light sources is as interpretable or usable as its colored incandescent counterpart. That assurance is especially important for pilots with marginal color vision such as those flying with color vision waivers.

Color Vision Screening for Pilots

The color vision requirements of the Code of Federal Regulations (CFRs) for aeromedical certification for first, second, and third class (Title 14 CFR Parts 67.103, .203, and .303; respectively) are the same, i.e., "Ability to perceive those colors necessary for the safe performance of airman duties." That requirement applies to both commercial and private pilots. The rationale is based on the international use of a color-coded signal system used by air traffic control specialists in control towers to direct aircraft in case of radio failure and other nonredundant uses of color such as position lighting on aircraft, runway/taxiway lighting, beacons, etc. (International Civil Aviation Organization, ICAO, 1988). Interpretation of the signal requires correct identification of three colors (red, green, and white) presented as steady, flashing, or alternating colors. Consequently, a color perception error could lead to an erroneous decision by the pilot and an increased risk of conflict with other aircraft. Individuals with normal color vision rarely confuse red, green, and white. However, it is well documented that individuals with color vision deficiencies frequently make errors involving those colors and many others.

Rensselaer Polytechnic Institute (RPI) Study

A study conducted by the Lighting Research Center at RPI (LRC RPI, 2011, p. 1) concluded:

> In general, LED signal lights resulted in improved color identification performance for color-normal observers. The results for protan and deutan observers were mixed with white and green LEDs resulting in improved identification performance relative to incandescent signals, but other colors showing no difference or worse identification performance for LEDs.

The RPI study examined several colors of incandescent lights and LEDs as did a follow-on study conducted at the FAA Civil Aerospace Medical Institute. However, the most pressing concern for the FAA Navigation Services, Lighting Systems Team (AJW-9142) is the implementation of LEDs into the PAPI lighting system; therefore, ensuring the usability of red and white LEDs is the first priority. There are four concerns to be addressed: 1) the RPI study had relatively few individuals with color vision deficiencies, with only nine individuals that failed the SLGT, 2) the distribution of individuals with color vision deficiencies was somewhat limited, according to the diagnostic tool, the Colour Assessment and Diagnosis (CAD) test, 3) a few individuals that passed the SLGT made a few errors on the red and white LEDs, the colors that are exclusively used in the precision approach path indicator (PAPI) light system, and 4) the RPI study involved other colors that are not inherent to the PAPI system (such as yellow, green, and blue).

Purpose

There were four main purposes of this study: 1) to expand the number of participants, especially those with color vision deficiencies, 2) to include a broader range of color vision abilities compared to previous studies, 3) to determine whether incorporating three shades of white LEDs in a cluster to compose the white light source, and likewise, three shades of red LEDs to compose the red light source may help to differentiate the red from white more successfully for individuals with color vision deficiencies, and 4) to limit the color set tested only to red and white.

METHODS

We obtained prior approval for all procedures and use of human subjects from the FAA Institutional Review Board. Participants provided informed consent prior to participation and subjects were free to withdraw from the project without consequence at any time.

Design and Protocol

Upon arriving at the FAA Mike Monroney Aeronautical Center in Oklahoma City, OK, the participants read and signed an informed consent form and were escorted as a group to an outside testing area for the Signal Light Gun Test. Researchers conducted the remaining tests inside the CAMI building. Participants performed some tests individually, such as the Dvorine Pseudoisochromatic plate test, the Colour Assessment and Diagnosis Test, and the Cone Contrast Test (CCT). The researchers administered other tests to subjects in small groups, such as the Signal Light Gun Test and a simulated precision approach path indicator task that was presented separately as incandescent and light emitting diodes (LEDs). For the tests presented to groups of subjects, each participant recorded his or her own responses on an individual answer sheet.

Visual Acuity

All subjects participated in a screening for both far and near visual acuity, with at least 20/30 in each eye with corrective lenses, if required, and using one or more of the following instruments:

1. Bausch and Lomb Orthorater - used for near and far visual acuity
2. Stereo Optical 5000 Vision Tester - used for near and far visual acuity
3. Titmus i400 Vision Tester - used for near and far visual acuity
4. U.S. Department of Transportation Federal Aviation Administration "Near Vision Acuity AAM Form 8500-1 (04-93)" was used for near visual acuity
5. Large Snellen "E" Test (model BC-11931, Bernell Corporation, South Bend, Indiana) was used for far visual acuity
6. Armed Forces Clinical Visual Acuity Test Form 3c - Chart used for far visual acuity

Color Vision Diagnoses

We examined participants' color vision using the Dvorine PIP, the CCT version 11, and CAD tests, and we based our discussions of diagnoses on the CAD. For general discussions, participants were classified as normal color vision (NCV) or as having color vision deficiencies (CVD) that are further classified by color vision type (protan—red weak, deutan—green weak, and tritan—blue weak).

Participants

A contractor recruited and paid 102 volunteers from the Oklahoma City commuting area. Sixteen of these subjects had participated in previous color vision studies over a three-year period; the researchers eliminated nine for missing data, resulting in usable data for 93 subjects (45 NCV, 48 CVD). We obtained the data reported in this study from subjects participating in a much larger study involving many more color vision tests. The sample contained 45 NCV, 27 deutans, 11 protans, two tritans, and eight subjects evidencing both red-green (RG) and yellow-blue (YB) deficiencies. NCV participants were 23 females and 22 males with a mean age of 23.9 years, SD of 3.5 years; CVD participants were 34 males and 14 females with a mean age of 23.8 years, and SD of 3.8 years. The minimum age was 18 and the maximum was 33 years. Age was restricted as a requirement of the validation of an air traffic control specialist applicant color vision screening test that was the primary research objective of the overall study. These data were collected as part of the larger study.

Apparatus and Materials

The Colour Assessment and Diagnosis (CAD) Test

In a joint FAA/Civil Aviation Authority (CAA) report, the level of color perception needed to perform the color-coded directions indicated by the PAPI lights was measured with the CAD test (Barbur, Rodriguez-Carmona, Evans, & Milburn, 2009) distributed by City Occupational, Ltd., London. The CAD test is a computerized color vision test that screens for normal color vision, quantifies loss of chromatic sensitivity, and classifies by type and degree of color vision deficiency. The full, definitive CAD test takes about 15 minutes to complete; however, unlike the Nagel anomaloscope, the traditional gold standard, diagnosis with the CAD test does not require an expert examiner to administer. The participant's task is to indicate the direction of movement of a colored square target on the dynamic checkerboard background via a response pad that employs a four-alternative, forced-choice procedure. Each of the four buttons corresponds to the four diagonal directions of movement.

In earlier versions of the test, the participant could repeat a trial by pressing the center button on the response pad (the direction of movement changed if repeated); however, in the most current version, repeating trials is only possible through the test administrator's keyboard. The very large number of trials prevents examinees from learning responses, which is possible on the limited trials of pseudoisochromatic plate tests. As an added benefit, the CAD test plots the individual's chromatic discrimination sensitivity in the CIE 1931 color space and provides both red/green and yellow/blue thresholds relative to the standard normal observer and reports those threshold values in standard normal units (SNUs) such that a threshold value of one indicates the normed value for the standard normal observer. No color naming is involved. The viewing distance from the 17-inch ViewSonic E70fSB CRT monitor is 140 cm (~55 inches). The illumination falling on the desktop in the testing room averaged about 10 to 15 lux.

The Dvorine Pseudo-Isochromatic Plates (PIP) 2nd Edition

The Dvorine PIP 2nd edition has 15 plates, with the pass criterion for normal color vision set at 13 correct of 15 plates established by the manufacturer. The FAA pass criterion for pilots is eight correct of 15 plates based on work by Mertens and

Milburn (1993). The Dvorine test booklet was displayed at a viewing distance of 61 and 76 cm (~2 to 2.5 ft) on the bookstand of the True Daylight Illuminator (both distributed by Richmond Products, Inc., Albuquerque, NM) equipped with a Verilux full spectrum (F15T8/VLX) lamp. The participants recorded their responses for each trial on an answer sheet.

Signal Light Gun Test (SLGT). Two different signal light guns were used in this study, the newer model, Model 901 (distributed by ATS Aerospace, Inc., Saint-Bruno, QC) and the older model, Crouse-Hinds Type W-1, which is no longer in production. This study will only present results involving the ATS Model 901 signal light gun, but we provide Table 1 and Figure 1 for reference for the two instruments. The primary differences between the two instruments: The ATS Model 901 is notably brighter, and the chromaticity of the green has moved away from the bluish green of the older instrument toward a yellowish green—note the change on the y-axis of Figure 1.

Table 1. 1931 CIE Coordinates and Cd/m^2 for the Crouse-Hinds and the ATS Aerospace Signal Light Guns

		CIE 1931 Coordinates		
	Color	Cd/m^2	x	y
Crouse-Hinds	Red	.460 E^2	.68	.27
Type W-1	Green	.712 E^2	.18	.35
(Old SLG)	White	1.311 E^2	.41	.36
ATS Aerospace	Red	.312 E^3	.69	.30
Model 901	Green	.613 E^3	.27	.57
(New SLG)	White	1.49 E^3	.41	.40

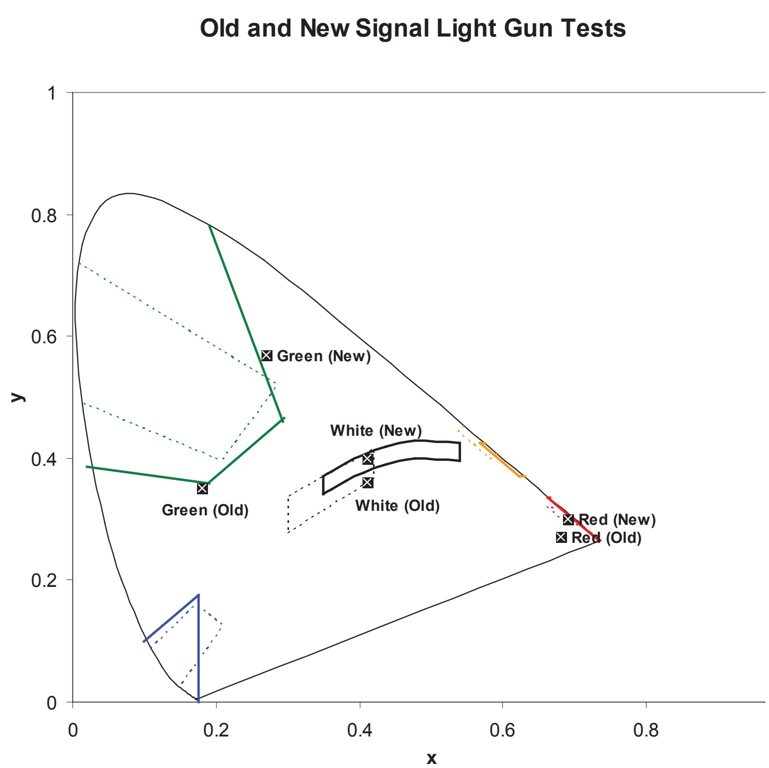

Figure 1. Comparison plots for the colored lights of the old and new signal light guns in CIE 1931 color space

The SLGT has a unique distinction, in that it is the actual instrument used by air traffic control specialists to communicate with pilots (see Figure 2), but it is also the same instrument used to determine whether a pilot receives a waiver for color vision. If a pilot applicant for a first- or second- class medical certificate fails an initial color vision screening test administered by an aviation medical examiner (AME), the applicant is required to take and pass an Operational Color Vision Test (OCVT) and a color vision Medical Flight Test (MFT). Applicants for a third-class medical certificate need only take and pass the OCVT. The OCVT has two components, the SLGT and demonstration of the ability to correctly read and identify colors on aeronautical charts, as outlined by FAA Order 8900.1 (FAA, 2014) for Issuance of a Medical Certificate and/or a SODA. Participants viewed the SLGT at two distances, a near distance of 1,000 feet and a simulated far distance of 1,500 feet. The far distance was simulated by placing a .30 neutral density Wratten filter in front of the signal light gun.

We first demonstrated the three aviation colors (red, green, and white) to acquaint the participants with the colors used in the signal light gun test. After demonstrating the colors, subjects were instructed on how to record their responses on the answer sheet, they recorded their subject identification number, the weather condition (cloudy or sunny), and were instructed not to disclose their responses aloud. This procedure created a time gap of at least 3 minutes between the demonstration of the colors and the beginning of the test. The time between trials was 2 minutes.

When FAA aviation safety inspectors give the SLGT to pilot-applicants, testing at the near distance is always first. However, as part of a separate study to determine whether continued testing at both distances is necessary, the ordering of the near and far distances alternated throughout the experimental trials. We presented the colors within each distance test site in the same order for all participants. In actual pilot applicant testing, examinees receive six trials at each distance with the three colors randomly ordered, and each color is presented at least once at each distance. With that criterion in mind, we defined a separate random order a priori for presentation at each distance for all participants. The group of participants took the SLGT as the first testing session after signing the consent form. Each participant was asked to circle, on the answer sheet provided, the name of the color presented.

Purposefully, we presented the colors twice at each distance in this experiment to control for the number of errors per color (Table 2). Normally, in actual pilot testing, the order of colors presented should be random and *may* present a color more than twice, or only once, at a location (e.g., Green, White, Green, Green, Red, and Red). Our decision to present each color twice at each location may have provided some predictability or expectancy to the test, thereby helping some individuals in our study—a benefit that may not occur in field testing. The pass criterion was zero errors among the 12 trials.

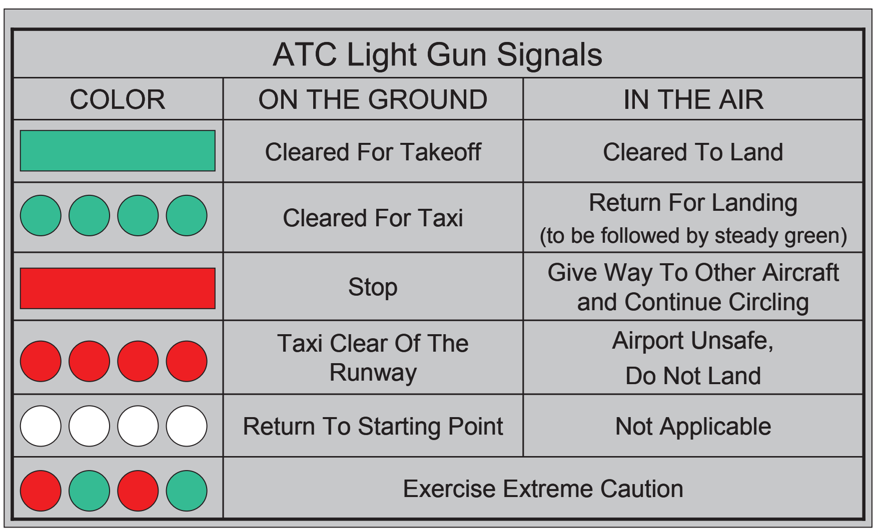

Figure 2. Air traffic control light gun signals and their meanings

Table 2. Colors presented for each trial with the SLG tests

Stimulus Number	Near (1,000 ft)		Far (1,500 ft)	
	Old SLG	New SLG	Old SLG	New SLG
1	White	Green	Red	Red
2	Green	White	Green	Green
3	Red	Green	Green	Green
4	White	White	White	White
5	Red	Red	Red	White
6	Green	Red	White	Red

According to a summary published in the 2011 Aerospace Medical Certification Statistical Handbook (Skaggs & Norris, 2013), the FAA currently has a total of 4,124 pilot special issuances for color vision distributed as follows: 1,543 first-class; 851 second-class; and 1,730 third-class pilots. This represents approximately 0.81% of first-class, 0.68% of second-class; and 0.65% of third-class pilots, notably a small percentage (0.71%) of all pilots.

PAPI Incandescent/LED Simulation Apparatus

A custom device developed at CAMI provided the emulation of the PAPI lighting system. This device presented a series of four horizontal lights to the participants. Using a pre-defined order, the experimenter presented the stimuli consisting of four lights of monochromatic and heterochromatic red and white 3-LED clusters or red and white incandescent lights (Figure 3). The stimuli were composed exclusively of either incandescent lights or LEDs, with the LEDs grouped into three monochromatic or three heterochromatic LEDs (Figure 3, Table 3).

Each of the incandescent lights consisted of 50W halogens with parabolic reflectors behind a 0.4375 inch (11.11 mm) aperture (Figure 4). The red incandescent light was created using filters constructed for current PAPI systems that were obtained from the FAA supply depot in Oklahoma City, OK. All of the LEDs were 5mm cylindrical LEDs. The LED lights were constructed with combinations of three LEDs (Table 3 and Table 4) behind 0.625 inch (15.88 mm) apertures (Figure 5).

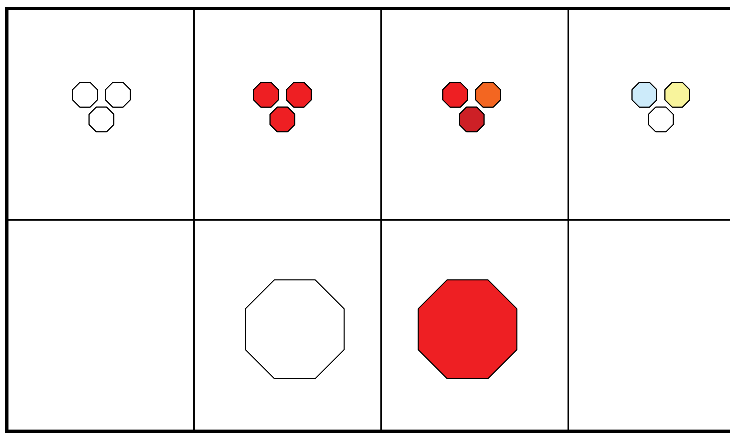

Figure 3. Sample box for PAPI task with LED clusters

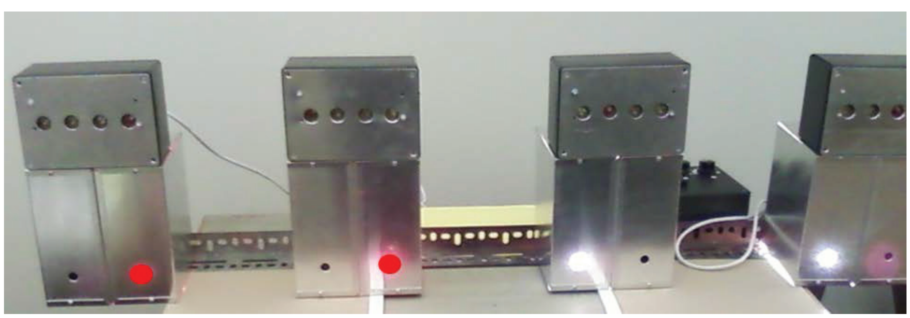

Figure 4. Photo of incandescent lamps illuminated in the experimental PAPI apparatus

Table 3. LED cluster combinations

Color	LED 1	LED 2	LED 3
Monochromatic White	5500K	5500K	5500K
Heterochromatic White	3000K	5500K	8000K
Monochromatic Red	642nm	642nm	642nm
Heterochromatic Red	625nm	642nm	660nm

Table 4. Characteristics of individual bulbs and LEDs

Color	Dominant Wavelength (nm)	Color Temp (K)	Illuminance (watts/m^2)	X Chromaticity Coordinate	Y Chromaticity Coordinate
LED Orange	625		3.5	0.71	0.29
LED Red	642		7.8	0.76	0.24
LED Red	660		5.3	0.77	0.23
LED White		3000	3.1	0.33	0.34
LED White		5500		0.32	0.33
LED White		8000	3.8	0.29	0.29
Incandescent White		2888	10.5	0.44	0.39
Incandescent Red		1000	2.5	0.67	0.33

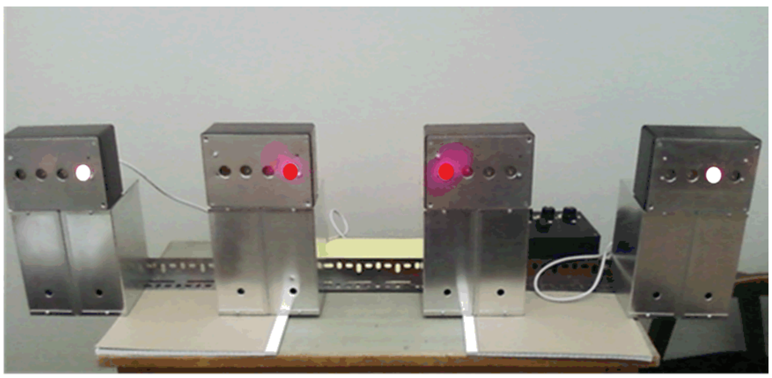

Figure 5. Photo of LEDs illuminated in the experimental PAPI apparatus

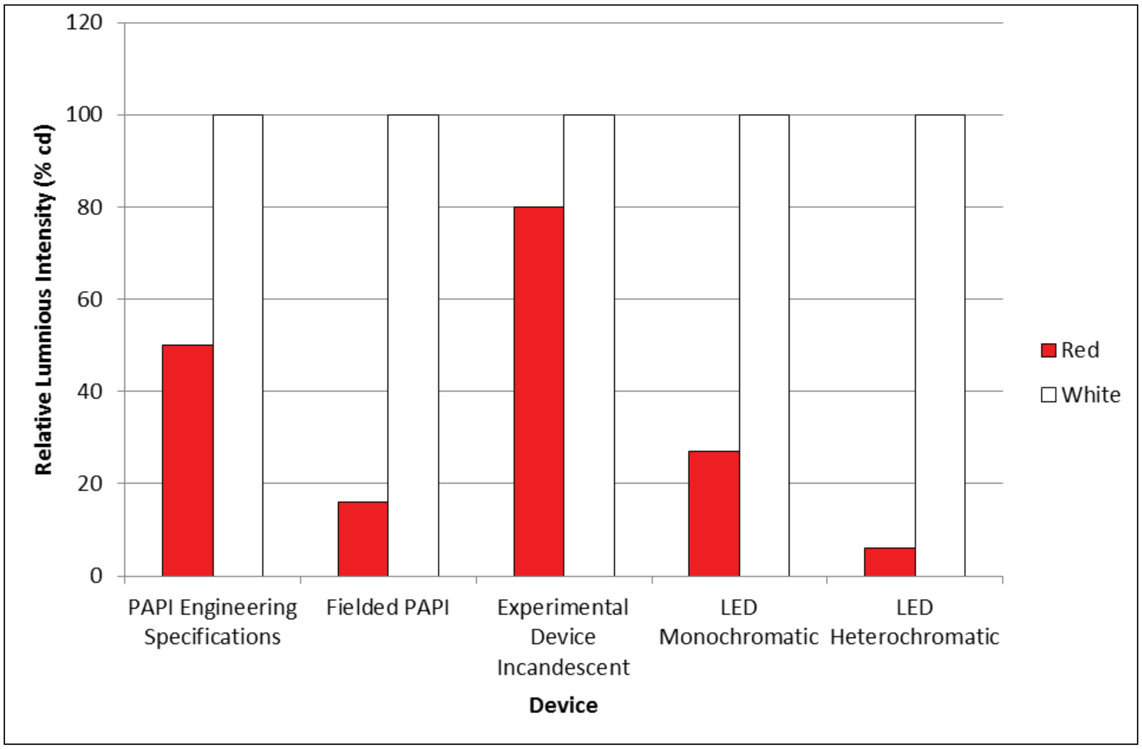

Figure 6. PAPI relative luminous intensities

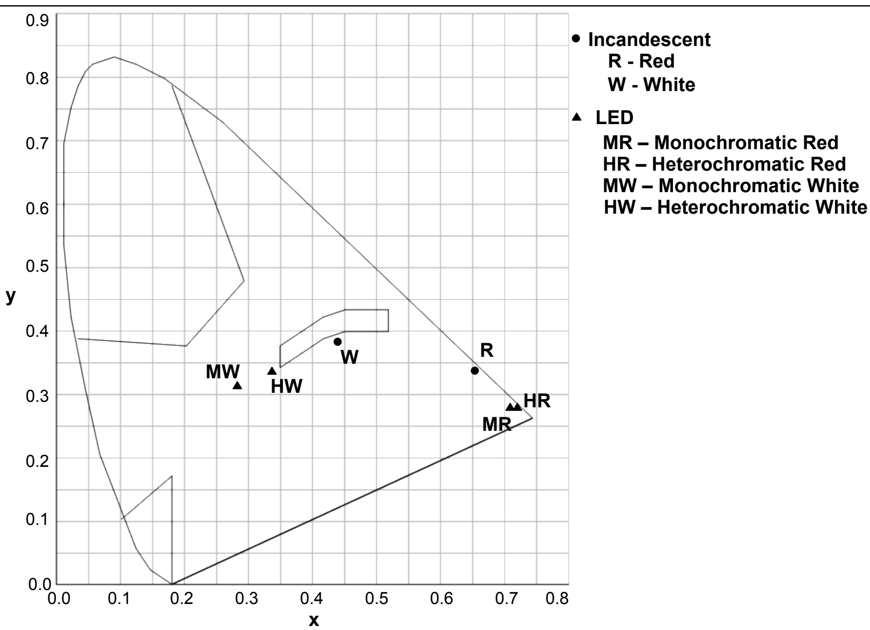

Figure 7. PAPI apparatus chromaticities plotted in CIE 1931 color space

The relative luminous intensities from the PAPI specifications (FAA, 2010a, 2010b, & 2011), one example of a fielded incandescent PAPI system (Astronics, 2014), and the light sources in our experimental device are represented in Figure 6.

The 625nm and 660nm red LEDs used colored filters in contrast to the other red and white LEDs, which had water clear lenses. The luminance of the 625nm and 660nm LEDs was less than the luminance of the other LEDs (Table 4). The PAPI apparatus chromaticities and color confusion lines for the protanopes, deuteranopes, and tritanopes are represented in Figures 7-10. The apparatus presented illuminances and chromaticities as indicated in Tables 4 and 5.

Participants viewed the apparatus from 215 feet (~ 65.5 meters), the length of a corridor. The positions of the lights in the boxes were not apparent from that distance, thus no positional cues were available to aid color identification.

At this distance, the degrees of arc for the incandescent lights was 0.0097 degrees (0.00017 radians or 35.00 arc seconds) and 0.0087 degrees (0.00015 radians or 31.49 arc seconds) for the LEDs, effectively creating a point source of light. The degrees of arc subtended were computed using the formula, V = 2 arctan(S / 2D), where V is the angle in radians, S the diameter of the aperture, and D the distance from the light source to the observer. These stimuli were similar to viewing actual PAPI lights at a distance of 3,936.9 feet (1.2 km) and 4,265.1 feet (1.3 km) from the airport respectively, given that the aperture for a light in a PAPI is approximately 8 inches (20.32 centimeters) in diameter.

LED arrays present additional challenges in determining distances where they become point sources. Previous analyses indicate that a three-LED array with Lambertian LEDs can be considered in the far-field when the array is a minimum of 8.5 times the distance of the array radius (Moreno & Sun, 2008; Moreno, Sun, & Ivanov, 2009). This places the beginning of the far-field for the LED arrays in our PAPI experimental device at approximately 85 mm. Thus, the light sources at 215 feet (65.5 m), including the LED arrays, in the PAPI experimental device were presented to the participants at a distance in the far-field.

Table 6 compares PAPI experimental device illuminances to equivalent illuminances in an actual PAPI.

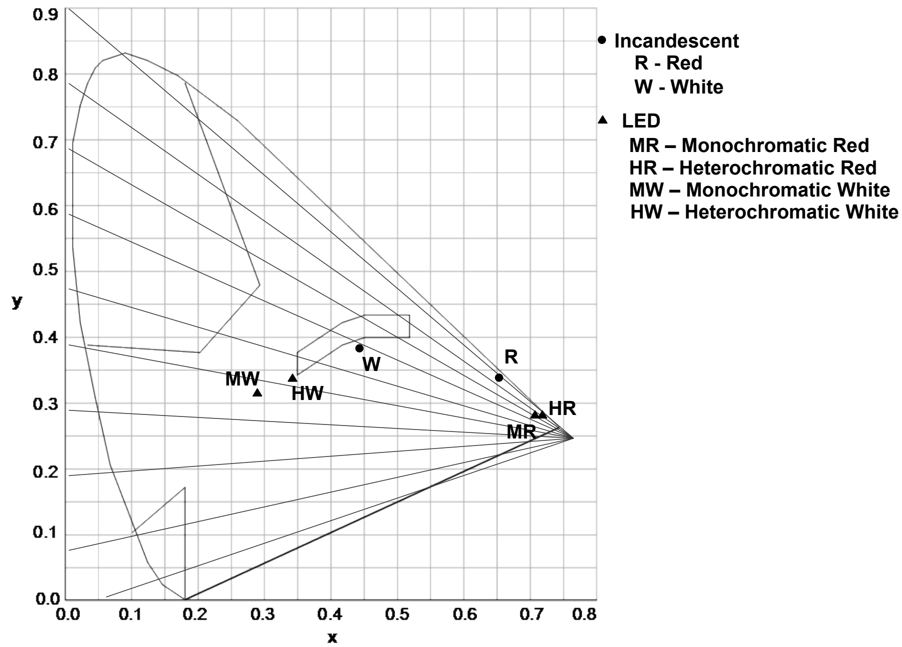

Figure 8. Color confusion lines for protanopes noted in CIE 1931 color space

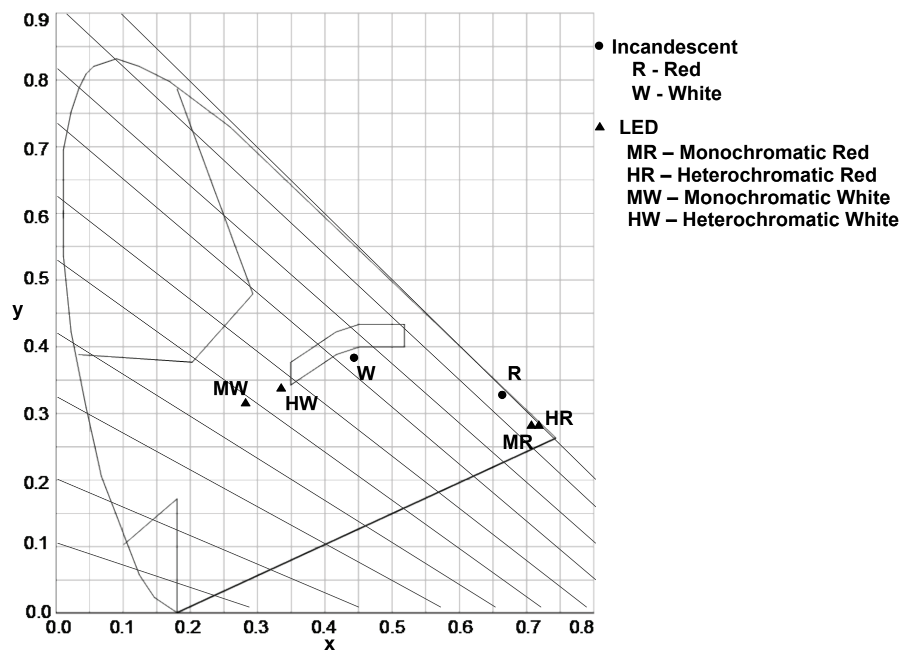

Figure 9. Color confusion lines for deuteranopes noted in CIE 1931 color space

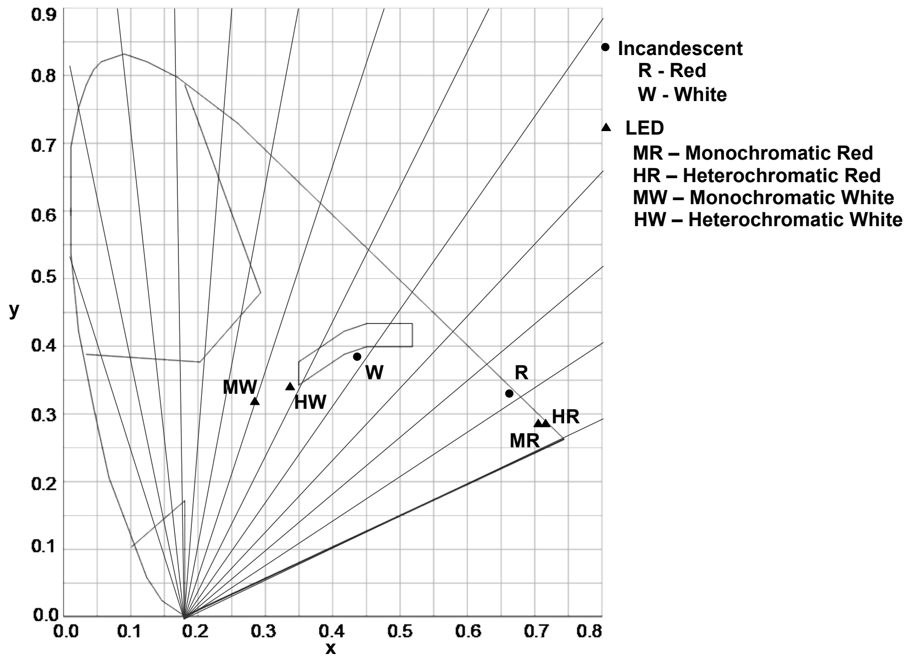

Figure 10. Color confusion lines for tritanopes noted in CIE 1931 color space

Table 5. Chromaticities and illuminance for 3-LED groupings and incandescent lamps

Color	x Chromaticity Coordinate	y Chromoticity Coordinate	Illuminance (watts/m^2)
LED Monochromatic White	0.28	0.31	3.2
LED Heterochromatic White	0.33	0.34	5.9
LED Monochromatic Red	0.72	0.28	2.8
LED Heterochromatic Red	0.72	0.28	1
Incandescent White	0.44	0.39	10.5
Incandescent Red	0.67	0.33	2.5

Table 6. Luminous intensities of experimental device

Color	Illuminance @ 0.0508 m (mlx)	Luminous Intensity @ 0.0508 m (cd)	Luminous Intensity @ 65.5 m (cd)	Equivalent to Fielded PAPI (m)
Monochromatic Red	115,000	0.29	0.0000001744	14897
Heterochromatic Red	34,000	0.09	0.0000000541	26740
Monochromatic White	420,000	1.08	0.0000006496	10917
Heterochromatic White	720,000	1.84	0.0000011068	8364
Incandescent Red	595,000	1.52	0.0000009143	6507
Incandescent White	733,000	1.88	0.0000011308	8274

The hallway was equipped with 13 overhead florescent office light fixtures each containing two 4 ft T35 tubes behind a diffuser (Sylvania 32 Watt Octron 3500K E40 lamps) producing 110 cd/m^2 at the LED/incandescent apparatus. The LED/incandescent apparatus was positioned between overhead light fixtures (Photograph 1). A substantial amount of sunlight was available from the large windows positioned approximately six feet behind the subjects.

Photograph 1. Photograph showing the LED/incandescent apparatus at the subject's viewing distance

PAPI Incandescent/LED simulation task

The experimenters manually selected the required combination of lights using four rotary switches prior to illumination. Participants recorded the color of each of the four lights for 40 trials by circling R for red, and W for white. Test administrators used a walkie-talkie to announce the trial number prior to presentation. There were five possible patterns (WWWW, WWWR, WWRR, WRRR, and RRRR) with each pattern presented an equal number of times.

Prior to testing, practice trials provided exposure to each of the lighting conditions: incandescent, monochromatic LEDs, and heterochromatic LEDs. Overall, participants responded to 12 practice lights (three trials composed of four lights each) and 160 total colored test lights (40 trials composed of a four-light PAPI simulation). Of those 160 total lights, there were an equal number of presentations of red and white lights (n=80) with half being presented as incandescent lights and the other half as LEDs.

The Cone Contrast Test (CCT, Rabin, Gooch, & Ivan, 2011) is a computer-based color vision test that indicates type (protan, deutan, or tritan), degree (mild to severe) of color vision deficiency, and quantifies normal color performance. Dr. Jeff Rabin of the United States Air Force's Operational Based Vision Assessment program developed the CCT as a CRT-based test, and it was recently marketed and distributed by Innova Systems, Inc., IL, as a Netbook version. Both versions were presented to subjects, with about half receiving the CRT-based test first and with a separate test between versions of the CCT; however, only the Netbook results (version 11) will be reported here because it is the commercially-available version. The CCT presents a sequence of colored, random letters visible to only one cone type at a time to provide a cone-specific numeric score.

The CCT examines each eye separately. The right eye was always tested first with the colored letters in red, green, and blue order (Rabin, Gooch, & Ivan, 2010). Twenty separate screen presentations, each comprised of a single, colored letter on a gray, achromatic background, were generated for each cone type (L, M, and S). In four increments with five letter-trials each, the amount of color saturation gradually fades to the gray background until the target color is hardly visible (Rabin, 1996). The CCT was administered in a darkened room at a viewing distance of 36 inches (91.44 cm). Participants were told that as the test proceeded the letters would become more desaturated and more difficult to see because the target letters become closer to the color of the background. The letters were displayed between the vertical and horizontal lines that appeared with each test stimulus. The participants were told that the test was short and fast-paced. The Netbook version allowed the participants to record their own responses, using the mouse to mark the letter on the response portion of the screen or the word *blank* if they did not see a letter.

RESULTS

Overview

In studies with multiple dependent variables (color, light source type, and monochromatic vs. heterochromatic clusters), in addition to between-groups variables (such as color vision ability), it is important to systematically approach the analyses to understand the results. To facilitate that, we divided the results into four major sections: I) Performance on the PAPI simulation, II) Predictive validity of a typical PIP screening test, III) Examination of the current waiver protocol, and IV) Predictive validity of computerized screening tests.

Section I will cover the performance analyses in the following order by examining: 1) the overall effects of light source, 2) differences between colors, 3) interaction effects of colors by light source, 4) interaction of light source by color vision classification, and 5) differences between monochromatic and heterochromatic clusters.

Section II will focus on the predictive validity of a typical initial color vision screening test to answer the second research question: whether the FAA's current color vision screening protocol adequately screens for the new LED technology. That series of analyses looks first at the Dvorine PIP test using two separate criterion, the Dvorine's manual designation of *normal* color vision (13 correct of 15), and the current FAA *pilot* passing criterion (8 correct of 15) to determine which criterion is better. The question then becomes, "If an applicant fails the initial screening but passes the waiver criterion (SLGT), will he or she be able to safely decode colored LEDs?" Those analyses will be covered in Section III. Finally, Section IV reports the predictive validity of two newer testing instruments, the CCT and the CAD, for performance on incandescent and LED light sources.

Section I. Examination of Performance by Color and Light Source

The PAPI simulation was composed of 40 trials, each containing four color-naming responses with an equal number of presentations for each light source type (incandescent vs. LED), color (red vs. white), and number of cluster compositions (monochromatic vs. heterochromatic LEDs). Percent correct performance was calculated for each subject. To answer the first research question: whether there is a performance difference between the incandescent and LED light sources, analyses were based on all trials for all subjects for each light source regardless of color vision status, light color, or cluster composition. Figure 11 indicates that participants performed slightly better on LEDs than incandescent lights. Performance for each light source, averaged over all participants, was very good—over 99% correct; note that the differences are exaggerated on the y-axis.

Participants were asked to identify the color (red or white) of incandescent and LED lights of the PAPI simulation. According to Figure 12, participants performed the worst on incandescent red lights. One deuteranope with additional yellow-blue deficiencies scored 31.58% correct on the red incandescent lights, and that individual did not pass the Signal Light Gun Test. About 96% of all subjects made fewer than two errors, resulting in an overall score between 95 and 100% correct (Table 7).

One important question is whether CVDs do worse on LEDs than on incandescent lights because of the planned conversion from incandescent to LEDs. Therefore, using a Repeated Measures GLM analysis with the between-groups defined as NCV or CVD, there were no significant effects for color or light source type and no interaction effects. Surprisingly, there were no between-groups effects—presumably because there were so many with CVDs that made no errors and a few individuals with NCV that made a few errors, hence blurring the separation between CVDs and NCV performance.

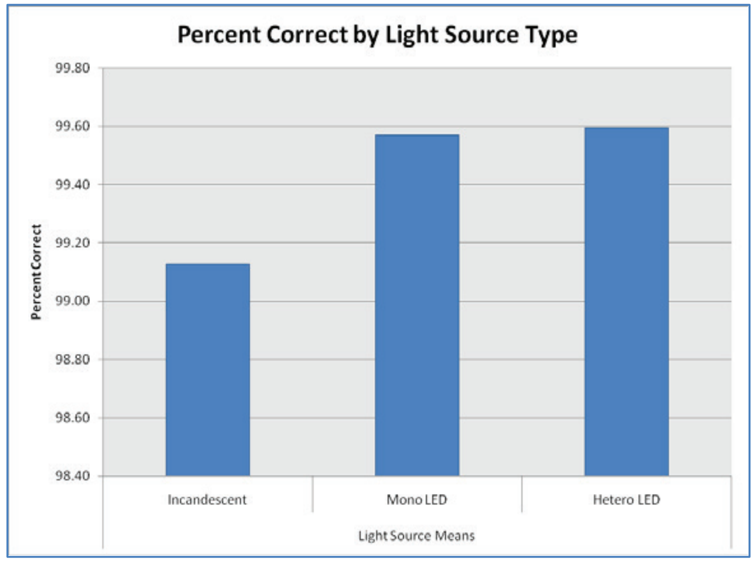

Figure 11. Percent correct by light source

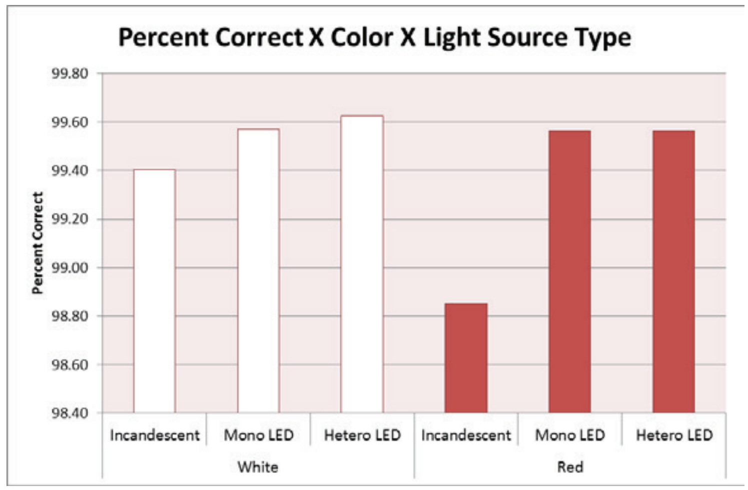

Figure 12. Percent correct by light source and color

Table 7. Frequency distribution of percent correct scores for all subjects, by color and light source type

	Frequency of Percent Correct					
	White			Red		
	INC	MC	HC	INC	MC	HC
31.58				1		
80.00						1
82.50	1					
85.00		1	1			
90.00	1	2	1		2	
91.89	1			1		
92.31				1		
92.50	1			1		
94.44						1
95.00	1	1	2	1	4	1
97.50	3			4		
100.00	85	89	89	84	87	89
Total	93	93	93	93	93	93

In a separate GLM analysis that defined groups by color vision type, Bonferroni multiple comparisons post hoc tests indicated that only the group with both Red-Green and Yellow-Blue deficiencies performed significantly different from all other groups, clearly displayed on Figure 13, which shows the performance on each light source type by color vision type classification (normal, protan, deutan, tritan, and red-green and yellow-blue deficiencies). Notice that performance was very good among all other color vision groups; unexpectedly, the protan group (n=11) performed perfectly on all light sources. Typically, protans perform worse on color identification tasks than NCV observers and deutans when luminance cues are absent (or masked by neutral density filters such as those used in the Farnsworth Lantern Test). Color naming tasks are especially difficult when differentiating red and white incandescent lights if reference colors are absent such as the PAPI simulator trials with four red or four white lights. Of the 40 PAPI task trials, 16 trials were composed of either four red or four white lights, providing ample opportunities for "guessing" strategies to fail.

Table 8 shows the numbers of subjects passing or failing with a pass criterion of 100% correct as a function of light source type (incandescent, LED MC, and LED HC), color, and CAD diagnosis.

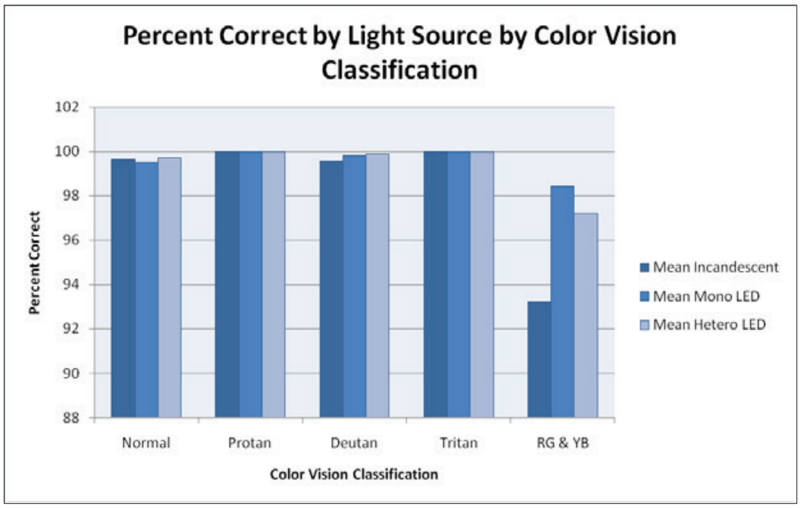

Figure 13. Performance on each light source type by color vision classification

Table 8. Frequency of passing (with 100% correct) and failing as a function of light source and color vision classification

Light Source	Normal N=45		Protan N=11		Deutan N=27		Tritan N=2		RG & YB N=8		ALL N=93	
	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail
Incandescent	40	5	11	0	23	4	2	0	4	4	80	13
Red	43	2	11	0	24	3	2	0	4	4	84	9
White	41	4	11	0	25	2	2	0	6	2	85	8
LED mono	42	3	11	0	26	1	2	0	5	3	86	7
Red	43	2	11	0	26	1	2	0	5	3	87	6
White	43	2	11	0	26	1	2	0	7	1	89	4
LED hetero	43	2	11	0	26	1	2	0	5	3	87	6
Red	44	1	11	0	26	1	2	0	6	2	89	4
White	43	2	11	0	27	0	2	0	6	2	89	4

Two groups were formed based on passing (with no errors) or failing the incandescent lights, to determine whether there were significant differences between groups on the LED lights. Table 9 provides the descriptive statistics for the two groups. Notice the following:

The lowest minimum score occurred for the LED hetero red with a score of 80 compared to the LED mono red counterpart with a score of 95 (both for the group that failed the incandescent lights).

For the failed group, the means are unexpectedly high for each of the comparisons—all above 96% correct.

Table 9. Descriptive statistics for performance on the LED tasks for two groups: those passing the incandescent task with zero errors and those failing the incandescent task

			Mean			95% Confidence Interval for Mean			
		N	Percent Correct	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
LED mono red	Fail	13	98.46	2.40	0.67	97.01	99.91	95	100
	Pass	80	99.75	1.57	0.18	99.40	100.10	90	100
	Total	93	99.57	1.75	0.18	99.21	99.93	90	100
LED mono white	Fail	13	97.69	4.84	1.34	94.77	100.62	85	100
	Pass	80	99.88	1.12	0.13	99.63	100.12	90	100
	Total	93	99.57	2.17	0.22	99.12	100.02	85	100
LED hetero white	Fail	13	97.31	4.84	1.34	94.38	100.23	85	100
	Pass	80	100.00	0.00	0.00	100.00	100.00	100	100
	Total	93	99.62	1.98	0.21	99.22	100.03	85	100
LED hetero red	Fail	13	96.88	5.98	1.66	93.27	100.49	80	100
	Pass	80	100.00	0.00	0.00	100.00	100.00	100	100
	Total	93	99.56	2.42	0.25	99.07	100.06	80	100

Significant between-groups differences were found in a one-way ANOVA (Table 10), using 100% correct on the incandescent lights as the pass criterion to form pass/fail outcome groups to test percent correct performance on the four LED compositions.

Table 11 provides the descriptive statistics for the six paired sample t-tests. There were no significant differences for any of the six planned paired comparison t-tests comparing the percent correct variable for each participant (Table 12).

- LED mono red vs. LED hetero red
- LED mono white vs. LED hetero white
- Incandescent white vs. LED mono white
- Incandescent red vs. LED mono red
- Incandescent white vs. LED hetero white
- Incandescent red vs. LED hetero red

These findings indicate that those that perform accurately on the incandescent lights will likely do so using LED lights, regardless of the composition of the light clusters.

Percent correct on incandescent red produced the largest SD (7.2) compared to the other lights that had SDs between 1.7 and 2.4 (see Table 11); however, the overall effect on the means was non-significant as shown on Table 12. As noted above, the large SD for incandescent red was caused by a single individual (a dichromat deutan with blue-yellow weaknesses) who scored 31.58% correct.

Table 10. ANOVA table for the analysis of percent correct for LED tasks with the between groups determined by passing or failing the incandescent lights task

		Sum of Squares	df	Mean Square	F	Sig.
LED mono red	Between Groups	18.57	1	18.57	6.39	.013
	Within Groups	264.23	91	2.9		
	Total	282.8	92			
LED mono white	Between Groups	53.28	1	53.28	12.77	.001
	Within Groups	379.52	91	4.17		
	Total	432.8	92			
LED hetero white	Between Groups	81.06	1	81.06	26.27	.000
	Within Groups	280.77	91	3.09		
	Total	361.83	92			
LED hetero red	Between Groups	108.83	1	108.83	23.07	.000
	Within Groups	429.35	91	4.72		
	Total	538.18	92			

Table 11. Percent correct descriptive statistics for the six paired samples

		Mean Percent Correct	N	Std. Deviation	Std. Error Mean
Pair 1	LED mono red	99.57	93	1.75	0.18
	LED hetero red	99.56	93	2.42	0.25
Pair 2	LED mono white	99.57	93	2.17	0.22
	LED hetero white	99.62	93	1.98	0.21
Pair 3	Incandescent white	99.40	93	2.42	0.25
	LED mono white	99.57	93	2.17	0.22
Pair 4	Incandescent red	98.85	93	7.22	0.75
	LED mono red	99.57	93	1.75	0.18
Pair 5	Incandescent white	99.40	93	2.42	0.25
	LED hetero white	99.62	93	1.98	0.21
Pair 6	Incandescent red	98.85	93	7.22	0.75
	LED hetero red	99.56	93	2.42	0.25

Table 12. Paired samples t-test results comparing mono vs. hetero and incandescent vs. LED by color, using the percent correct variable

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Pair 1	LED mono red vs. LED hetero red	0.01	2.46	0.25	-0.50	0.51	0.02	92	0.98
Pair 2	LED mono white vs. LED hetero white	-0.05	1.38	0.14	-0.34	0.23	-0.38	92	0.71
Pair 3	Incandescent white vs. LED mono white	-0.17	2.53	0.26	-0.69	0.35	-0.64	92	0.52
Pair 4	Incandescent red vs. LED mono red	-0.72	6.82	0.71	-2.12	0.69	-1.02	92	0.31
Pair 5	Incandescent white vs. LED hetero white	-0.22	2.31	0.24	-0.70	0.25	-0.92	92	0.36
Pair 6	Incandescent red vs. LED hetero red	-0.71	5.16	0.54	-1.78	0.35	-1.33	92	0.19

Section II. Validity of the Dvorine PIP for Predicting Performance on LED and Incandescent Red and White Lights

Most pilots satisfy the color vision screening requirement by taking an FAA-approved PIP test such as the Dvorine. So, the question is: How well does the Dvorine PIP screen for ability to identify the red and white lights of the PAPI system? The current PAPI system is composed of incandescent lights, unfiltered to produce white and with a red glass filter to create a red light. The baseline for measuring performance improvements or decrements for the LEDs should be compared with performance on the incandescent lights.

Fifty-nine subjects passed the Dvorine (using the manufacturer's manual criterion of 13 of 15 correct), and 34 subjects failed. Using that Dvorine pass/fail outcome as a grouping variable and the percent correct score for each light source as the dependent variable in a Repeated Measures GLM, there were no significant main effects due to color or light source, no significant interaction effects, and no significant between-groups effect. To better understand that last non-significant finding, a crosstabulation table of pass/fail on the Dvorine was constructed for each light source. The pass criterion for each light source was based on 100% correct to pass.

CAD analyses indicated that seven (including the two individuals with a tritan deficiency) of the 59 subjects that passed the Dvorine had some degree of yellow-blue CVD, which the Dvorine does not measure.

The Kappa (agreement) scores were very low between the Dvorine and the lights, all less than .07, showing essentially no agreement between pass/fail outcomes. Notice (Tables 13 and 14) the large number of individuals failing the Dvorine (Manual criterion) that passed each of the light sources indicated by false positive rates ~ 35%. In a similar comparison, but using the FAA's pilot criterion for pass/fail on the Dvorine (pass with 8 or more correct), the sensitivity values do not look much different. The root cause of the low agreement (Kappa) scores for the comparisons can be isolated to the individuals that fail the Dvorine (Manual and Pilot criteria) and pass the light tasks; however, from a safety standpoint, those (false positives) are *not* the individuals of major concern. Instead, the individuals that *passed* the Dvorine and *failed* the lights tasks (misses or false negatives) are of concern because they would be cleared based on the initial color vision screening test and subsequently be unable to differentiate the PAPI light colors. There were more participants that passed the screening but failed on the incandescent lights (~12%) compared to "misses" for the LED lights (~6%); but screening test *misses* are never a good thing. Based on these results, it would appear that the screening test is not adequately differentiating those that do not possess the ability to accomplish the job task. This laboratory apparatus was meant to be a simplistic representation of PAPI lights, isolating only a few perceptual aspects that may be involved in actual PAPI systems. Pilots are probably using many other visual cues to compensate or complement the PAPI system such as the runway "sight picture" to judge angle of descent (Mertens, 1979).

Section III. Performance of Those Waivered by the SLGT

In this sample, 33 participants failed the Dvorine, and 11 of those passed the SLGT. If the initial screening decision were based on the Dvorine, approximately 33% of pilot applicants that failed the Dvorine would benefit from secondary screening using the SLGT and would have qualified for waivers. Those 11 individuals had color vision classifications as follows: one NCV, three protans, five deutans, and two that were both RG and YB weak.

Table 15 provokes two important questions: How did those 11 "waivered" individuals perform on the color-coded lights tasks and why did those six subjects pass the Dvorine (manual criterion) yet fail the new signal light gun test?

Table 13. General representation of a cross-tabulation table for calculation of sensitivity, specificity, false positive, and false negative

		Criterion Measure/Reference Test		
		Pass	Fail	Total
Screening Test	Pass	*a*	*b*	*a* + *b*
	Fail	*c*	*d*	*c* + *d*
	Total	*a* + *c*	*b* + *d*	N

Sensitivity = $d / (b + d)$ and specificity = $a / (a + c)$
False Alarm (false positive) = $c / (a + c)$
Miss (false negative) = $b / (a + b)$

Table 14. Sensitivity, specificity, false positive, and false negative rates for the Dvorine PIP using two pass/fail criteria

		Sensitivity	Specificity	False Positive (False Alarm)	False Negative (Miss)
Dvorine with Manual Criterion	Incandescent	46.1%	65.0%	35.0%	11.8%
	LED Mono	42.8%	63.9%	36.0%	6.7%
	LED Hetero	50%	64.3%	35.6%	5.1%
Dvorine with Pilot Criterion	Incandescent	38.4%	71.2%	28.7%	12.3%
	LED Mono	42.8%	70.9%	29.0%	6.1%
	LED Hetero	50.0%	71.2%	28.7%	4.6%

Table 15. Crosstabulation of the Dvorine (Manual Criterion) Pass/Fail by the New Signal Light Gun Pass/Fail

		New Signal Light Gun		
		Fail	Pass	Total
Manual Dvorine	Fail	21	11	32
	Pass	6	51	57
Total		27	62	89

Note: Missing data for 4 subjects because the battery failed during testing

Table 16 shows that, in general, the 11 "waivered" individuals were able to distinguish the red and white LEDs; and, that performance on the incandescent lights was somewhat worse.

The second question concerns the six individuals, diagnosed as follows: two NCV, one deutan, one tritan, and two with RG & YB weaknesses that passed the Dvorine (manual criterion) but failed the new signal light gun test. If these individuals had been actual pilot applicants, their score on the Dvorine would have been sufficient to pass the FAA's color vision screening process—and they would not have had to take and pass the SLGT. So, it is only this research data that affords us the opportunity to see the inter-relationship between the initial screening test (the Dvorine) and the FAA's secondary screening test (the SLGT) when the participant *passes* the initial test. Upon reviewing their performance on those two tests, we discovered that the two NCV participants (0633 and 0660) made one error each on the signal light gun test and no errors on the Dvorine. The CAD diagnosed subjects 0645 and 0650 with a deutan deficiency and noted that a "yellow-blue acquired deficiency [is] likely."

At the time this study was conducted, the CAD did not have a yellow-blue cut-point for aeromedical certification. Subject 0650's threshold for YB was 8.89 and subject 0645's YB threshold was 4.52, both well outside the normal limits. Both subjects missed only plate 15 of the Dvorine. Subject 0638 made no errors on the Dvorine and had CAD threshold values of 1.69 and 1.84 for RG and YB, respectively. The CAD diagnosed her with normal RG color vision and deficient YB color vision. She had one error on the SLGT. Subject 0611 made one error on the SLGT and had a CAD RG and YB thresholds of 2.40 and 1.34, respectively, resulting in diagnoses of deutan deficiency and normal YB color vision. Her only error on the Dvorine was on plate 15.

Table 16. CAD Diagnosis by Pass/Fail performance on light source type and color for the 11 "waivered"

		CAD Type Diagnosis				
		Normal Count	Protan Count	Deutan Count	Tritan Count	RG & YB Count
INC White	Fail	1	0	2	0	1
	Pass	0	3	3	0	1
INC Red	Fail	1	0	1	0	1
	Pass	0	3	4	0	1
LED Mono White	Fail	0	0	1	0	1
	Pass	1	3	4	0	1
LED Mono Red	Fail	0	0	1	0	1
	Pass	1	3	4	0	1
LED Hetero White	Fail	0	0	0	0	1
	Pass	1	3	5	0	1
LED Hetero Red	Fail	0	0	1	0	1
	Pass	1	3	4	0	1

Table 17 was constructed showing the CAD diagnosis of those *passing* the SLGT and their performance on the PAPI simulation task designated as the following subtests: INC, Mono and Hetero LEDs. Kappa scores were all non-significant and essentially zero, primarily because of the large number of individuals that passed the PAPI simulation lights but failed the SLGT.

Table 17. The pass/fail outcome on the PAPI simulator lights by CAD type diagnosis for only those individuals that **passed** the SLGT

		CAD Type Diagnosis				
		Normal Count	Protan Count	Deutan Count	Tritan Count	RG & YB Count
INC White	Fail	3	0	2	0	1
	Pass	37	3	11	1	4
INC Red	Fail	2	0	2	0	2
	Pass	38	3	11	1	3
LED Mono White	Fail	2	0	1	0	1
	Pass	38	3	12	1	4
LED Mono Red	Fail	2	0	1	0	1
	Pass	38	3	12	1	4
LED Hetero White	Fail	1	0	0	0	1
	Pass	39	3	13	1	4
LED Hetero Red	Fail	1	0	1	0	1
	Pass	39	3	12	1	4

According to Table 18, about 89 to 93% of those *failing* the SLGT *passed* the subtests of the PAPI simulation test. If there were only a few individuals that performed differently, it would be easier to blame the incongruent findings on factors unrelated to the tests, but the overwhelming performance improvement is somewhat disconcerting because one would expect similar performance on the incandescent lights and could attribute the improved performance on the LEDs as being related to the differences in spectral characteristics (e.g., the narrower band compared to that of incandescent lights). Speculation as to the cause of the disagreement (between the SLGT and the PAPI simulation) test outcomes includes differences resulting from disparate viewing distances, color saturations, number of color choices, ambient lighting conditions, background clutter, and even personal comfort during testing. Participants may have been less attentive outside vs. inside because of the warm Oklahoma summer weather (~8:30 to 9:00 A.M.) compared to inside with air conditioning. One major difference between the incandescent PAPI simulation and the SLGT was that the latter required the identification of the color green, in addition to the red and white lights. Upon closer examination of the errors on the SLGT, we discovered that 19 of the 27 participants that failed the SLGT did so because of errors on green lights. This information shines light on the cause of the disparate pass/fail outcomes between the SLGT and the PAPI lights because deutans are more likely to confuse green with white lights and the incandescent and LED lights of the PAPI system only require red and white lights, a combination easier for deutans to discern, hence allowing more deutans to pass the simulated PAPI task more often than the SLGT. Consequently, the SLGT is not necessarily a good indicator of how an individual will perform on PAPI red and white lights, but it is valuable as a screening test for other aviation tasks such as discerning navigational lights with its red light on the left wing tip and green light on the right to note an aircraft's heading and direction of travel. Acknowledging the greater difficulty that deutan-type deficient have with green and white lights, the signal light gun is still in use; and, according to Figure 2, a pilot must be able to tell the difference between green and white lights to determine the meaning of the air traffic controller's directions (e.g., four flashes of white light vs. four flashes of green light meaning "return to starting point" or "cleared for taxi," respectively). Clearly, all subjects performed well on the simulated PAPI tasks, regardless of whether the lights were incandescent or LEDs. Furthermore, keep in mind that although interpretation of PAPI lights is considered a safety-critical task, it is not the *only* color-coded safety-critical task nor the most difficult, so even though performance on the SLGT is not a perfect match to the PAPI system, its value as a screening test may be more related to other color-coded tasks, especially the pilot's ability to interpret the emergency use of the signal light gun in cases of radio failure and tasks that involve greater color perception such as cockpit displays, maps, and charts.

Table 18. The pass/fail outcome on the PAPI simulator lights by CAD type diagnosis for only those individuals that **failed** the SLGT

		CAD Type Diagnosis				
		Normal Count	Protan Count	Deutan Count	Tritan Count	RG & YB Count
INC White	Fail	0	0	0	0	1
	Pass	2	7	14	1	2
INC Red	Fail	0	0	1	0	2
	Pass	2	7	13	1	1
LED Mono White	Fail	0	0	0	0	0
	Pass	2	7	14	1	3
LED Mono Red	Fail	0	0	0	0	2
	Pass	2	7	14	1	1
LED Hetero White	Fail	0	0	0	0	1
	Pass	2	7	14	1	2
LED Hetero Red	Fail	0	0	0	0	1
	Pass	2	7	14	1	2

Section IV. Predictive Validity of Several Screening Tests for the PAPI Simulation Tasks

Three summary tables (19, 20, & 21) were constructed showing the pass/fail crosstabulation between the CAD certification, the CCT version 11, the Dvorine, SLGT and each of the PAPI simulation tasks. The aeromedical certification cut-point established by CAD developers (Barbur et al., 2009), which was based on laboratory performance using colored incandescent lights, was purposefully set slightly outside the boundaries of normal color vision to include those that can identify the colored lights, yet do not possess normal color vision for the purpose of creating a more lenient pass criterion. However, according to these data, that criterion still appears to be more stringent than supported by the CCT, the Dvorine (manual or pilot criterion), and the secondary screening test used to issue waivers—the Signal Light Gun Test. Eighty-nine participants completed all four of those tests, with the following number of individuals passing: CAD certification, 51; Netbook CCT, 56; Dvorine (manual criterion), 59; Dvorine (pilot criterion), 65; and the SLGT, 62. According to Tables 18, 19, and 20, the CCT showed almost the same pattern of pass/fail performance on the lights tasks as the Dvorine, with the greatest disparate group being those that failed the screening test but passed the lights tasks. Once again, from a safety standpoint, the small group of individuals that *passed* the screening test and *failed* the lights tasks (incandescent and LEDs) are the ones of concern; however, performance on the proposed replacement (LEDs) was better than on the existing incandescent technology. Essentially all of the examined screening tests failed some individuals that possessed sufficient color perception to identify the colored lights presented in this laboratory simulation task.

Table 19. Composite crosstabulation tables showing pass/fail outcome on the incandescent PAPI lights with screening and practical tests

		Incandescent PAPI	
		Fail Count	Pass Count
CAD Certification	Fail	7	35
	Pass	6	45
CCT	Fail	6	31
	Pass	7	49
Dvorine (Manual Criterion)	Fail	6	28
	Pass	7	52
Dvorine (Pilot Criterion)	Fail	5	23
	Pass	8	57
New Signal Light Gun Test	Fail	3	24
	Pass	9	53
Hetero LED PAPI	Fail	6	0
	Pass	7	80
Mono LED PAPI	Fail	5	2
	Pass	8	78

Table 20. Composite crosstabulation tables showing pass/fail outcome on the Mono LED PAPI lights with screening and practical tests

		Mono LED PAPI	
		Fail Count	Pass Count
CAD Certification	Fail	4	38
	Pass	3	48
CCT	Fail	3	34
	Pass	4	52
Manual Dvorine	Fail	3	31
	Pass	4	55
Pilot Dvorine	Fail	3	25
	Pass	4	61
New Signal Light Gun	Fail	2	25
	Pass	5	57
Incandescent PAPI	Fail	5	8
	Pass	2	78
Hetero LED PAPI	Fail	4	2
	Pass	3	84

CONCLUSIONS

In general, all of the screening tests reported in this study—CAD certification, Dvorine (pilot or manual criteria), the CCT, and the SLGT—were very conservative predictors of performance on the red and white incandescent and the LED PAPI simulations because a high percentage of our sample made no errors on the PAPI simulations but failed the screening test, thereby producing a high false alarm rate, also called false positive (e.g., CAD certification false alarm=43.6% for the HC LED PAPI, reference Table 21). The close similarity of performance on the LED PAPI simulations (mono vs. hetero LEDs and red vs. white LEDs) gives some evidence that the ability level of color vision required to interpret non-redundant PAPI color coding using LEDs is relatively low and that passing a high-quality color vision screening test (such as the Dvorine, CCT, or CAD certification test) will almost assure that the examinee is capable of distinguishing red and white LED lights of a similar chromaticity to those used in this study; however, those same screening tests are likely to produce only moderate sensitivity and specificity rates because of the high false alarm rate. From a safety standpoint, these screening test results are desirable—because they miss very few, meaning that rarely do they pass individuals that cannot perform the necessary color tasks.

Given the increased usage of color in cockpit displays, it is highly likely that, for pilots, the LED PAPI task (although a critical task) is not the most demanding color discrimination task that pilots will face; and therefore, the usefulness of these screening tests for determining whether the pilot possesses the requisite color vision ability for other tasks, including tasks involving multi-colored LEDs, should not be based on the validity of the tests for predicting performance on this red vs. white LED task. Furthermore, it is not known whether these PAPI simulation test results are better or worse than performance likely on-the-job. A few of the factors that might make the simulation results better might be effects not represented in a laboratory research environment such as increased time pressures, negative consequences of wrong color identifications, adverse atmospheric conditions, and interferences and interruptions, both cognitive and auditory, just to mention a few. Conversely, the laboratory experiment was not necessarily a "sterile" environment devoid of interferences and interruptions because the PAPI simulation task was conducted using the length of a long hallway that occasionally experienced foot traffic between test trials; however, between trial interruptions is much different from an interruption that occurs on approach. Although these findings seem to indicate that LEDs, whether hetero or monochromatic, are

Table 21. Composite crosstabulation tables showing pass/fail outcomes on the Hetero LED PAPI lights with screening and practical tests

		Hetero LED PAPI	
		Fail Count	Pass Count
CAD Certification	Fail	4	38
	Pass	2	49
CCT	Fail	3	34
	Pass	3	53
Manual Dvorine	Fail	3	31
	Pass	3	56
Pilot Dvorine	Fail	3	25
	Pass	3	62
New Signal Light Gun	Fail	2	25
	Pass	3	59
Incandescent PAPI	Fail	6	7
	Pass	0	80
Mono LED PAPI	Fail	4	3
	Pass	2	84

suitable for incandescent replacements, there are some critical questions remaining that we will address in a future paper based on the follow-on research.

Follow-On Research

We used a StellarNet LT-16 spectrometer with an integrating sphere to make the measurements reported in Figure 6. The substantial differences in relative luminance intensities between the reds and whites of each condition present an extreme case of what one might find in a fielded system. For example, one particular fielded PAPI system presents a six to seven-fold increase in relative intensity from the red and white lights (Astronics, 2014). Notably, we adjusted the brightness differences to a point where they *were not apparent* to the participants in the preliminary prototype evaluation, which included a small sample of NCV and CVD subjects. However, PAPI systems could potentially have the relative luminous intensity differences outlined in the engineering specifications of 2:1 with the white being twice the intensity of the red. Because of this, we are uncomfortable relying on the current results to make recommendations for mono vs. hetero clusters of LEDs or recommending non-redundant usage of (red and white) color coded LEDs for individuals with marginal color vision until we can explore the effects that differing luminous intensities may have on both LED and incandescent signal light color identification. Therefore, we designed a follow-on experiment that is very similar but the important difference being that we can tightly control the relative luminance intensities for all light stimuli. We have purchased a Photo Research 740 spectroradiometer for that purpose. We found that the LED measurements made with the StellarNet LT-16 spectrometer integrating sphere were comparable to those taken with the Photo Research 740. Furthermore, we will explore the effect of presenting white LEDs at twice the brightness of red, a condition that matches the current engineering specifications for incandescent PAPI systems. Additionally, we presented the two colors at the same relative luminance intensity to evaluate performance in the event that some future PAPI system presents the red and white signals at equal luminous intensities (Gildea & Milburn, 2013). Those results will appear in a separate, forthcoming report.

Based on the results reported here, we can safely say that performance on red and white lights, whether LEDs or incandescent, is very similar for those passing the screening tests, and those individuals are likely to experience few problems identifying the colored lights of the PAPI system. The two points in question are whether the relative luminance intensities between 1) the reds vs. whites, and 2) the monochromatic vs. the heterochromatic clusters provided a redundant cue that assisted CVD participants. The experiment by Bullough et al. (2012) accurately controlled the relative luminance intensities, and they reported improved performance for LEDs by NCV observers and mixed results for CVDs ranging from worse identification to no difference relative to incandescent lights. However, there were several design differences between the present experiment and Bullough's experiment that we believe account for the performance differences reported here. For example, the Bullough experiment involved many different colors, whereas our experiment isolated only the red and white colors of the PAPI system, thereby improving the probability of a correct response. The Bullough experiment did not explore the effect of monochromatic vs. heterochromatic clusters of LEDs. The follow-on experiment we are planning should clear up these discrepancies and allow the FAA to implement lighting systems and screening tests that are both reliable and effective to as wide a pilot population as possible.

REFERENCES

Astronics DME Corporation (2014). *PAPI: Precision approach path indictor*. http://www.astronics.com/_images/airfield-lighting/AstronicsDME%20PAPI%20E-spec%20Data%20Sheet%20050311.pdf Accessed January 28, 2014.

Barbur, J., Rodriguez-Carmona, M., Evans, S. & Milburn, N. (2009). *Minimum color vision requirements for professional flight crew, Part III: Recommendations for new color vision standards.* (Report No. DOT/FAA/AM-09/11). Washington, DC: Federal Aviation Administration Office of Aerospace Medicine. Online at: http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/

Barbur, J.L., Rodriguez-Carmona, M., & Harlow, A.J. (2006). *Establishing the statistical limits of 'normal' chromatic sensitivity*. CIE Expert Symposium, CIE Proceedings 75 Years of the Standard Colorimetric Observer.

Bullough, J.D., Skinner, N.P., Bierman, A., Milburn, N.J., Taranta, R.T., Narendran N., & Gallagher D.W. (2012). *Nonincandescent source aviation signal light colors*. (Report No. DOT/FAA/TC-TN12/61). Washington, DC: Federal Aviation Administration.

Code of Federal Regulations (CFR), Title 14, Part 67. Volume 5, Airman Certification. Chapter 8, Conduct a special medical test, Section 1 Issuance of a Medical Certificate and/or a SODA, or LOE. Washington, DC: U.S. http://www.faa.gov/documentLibrary/media/Order/8900.1.pdf Accessed March 19, 2014.

Commission Internationale de L'Eclairage (2001). *International recommendations for colour vision requirements for transport.* (Technical Report No. CIE 143). Vienna: CIE.

FAA (1988). Federal Aviation Regulations Part 23, Sec 23.1397. Washington, DC: 1988.

FAA (2009). *Specification for runway and taxiway light fixtures* (AC 150/5345-46).

FAA (2010a). *Design and installation details for airports visual aids* (AC 150/5340-30).

FAA (2010b). *Airport lighting equipment certification program* (AC 150/5345-53).

FAA (2011). *Precision approach path indicator (PAPI Systems).* Advisory Circular No. 150/5345-28G. http://www.faa.gov/documentLibrary/media/Advisory_Circular/150_5345_28g.pdf. Accessed January 28, 2014.

Gildea, K.M., & Milburn, M. (2013). *Open-source products for a lighting experiment device.* Behavior Research Methods. Retrieved from http://link.springer.com/article/10.3758%2Fs13428-013-0423-1International Civil Aviation Organization (ICAO) (1988). Technical airworthiness manual (2nd ed. Document 9051-AN/896), Amendment 10. Montreal, Quebec, Canada.

Lighting Research Center at Rensselaer Polytechnic Institute (LRC RPI) (2011). *Color identification of LED aviation signal lights.* Project Sheet: http://www.lrc.rpi.edu/resources/newsroom/pdf/2011/LEDColorID8511.pdf

Mertens, H.W. (1979). *Runway image as a cue for judgment of approach angle.* (Report No. DOT/FAA/AM-79/25). Washington, DC: Federal Aviation Administration Office of Aviation Medicine. Online at: http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/

Mertens, H.W. & Milburn, N.J. (1993). *Validity of FAA-approved color vision tests for class II and class III aeromedical screening.* (Report No. DOT/FAA/AM-93/16). Washington, DC: Federal Aviation Administration Office of Aviation Medicine. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a271739.pdf

Moreno, I. & Sun, C. (2008). *LED array: Where does the far-field begin?* In I.T. Ferguson, T. Taguchi, I.E. Ashdown, & S. Park (Eds.) Proceedings for SPIE Vol. 7058. Retrieved from http://fisica.uaz.edu.mx/~imoreno/Publicaciones/SPIE2008a.pdf

Moreno, I., Sun, C., & Ivanov, R. (2009). *Far-field condition for light-emitting diode arrays*. App Opt*, 48,* 1190-1197.

Rabin, J. (1996). *Cone specific measures of human color vision*. Invest Ophthalmol Vis Sci. 1996; 37:2771-74.

Rabin, J., Gooch, J., & Ivan, D. (2010). *Rapid quantification of color vision: The cone contrast test.* Invest Ophthalmol Vis Sci. 2011 Feb 9;52(2):816-20. Retrieved from http://www.iovs.org/content/52/2/816.full.pdf

Rodriguez-Carmona, M. (2006). *Variability of chromatic sensitivity: Fundamental studies and clinical applications.* Unpublished Doctoral thesis, City University, London, United Kingdom.

Skaggs, V. & Norris, A. (2013). *2012 Aerospace medical certification statistical handbook.* (Report No. DOT/FAA/AM-13/25). Washington, DC: Federal Aviation Administration Office of Aerospace Medicine. Online at: http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/